The Fourth International Conference on Computational Systems Biology (ISB2010)
Suzhou, China, September 9–11, 2010

Significant Regulatory Networks from Goto-Kakizaki Rat Liver Microarray Data during Diabetic Progression

Huarong Zhou[1] Shigeru Saito[2,3] Guanying Piao[4] Rongkuan Hu[4]
Jiguang Wang[1] Zhi-Ping Liu[1] Katsuhisa Horimoto[2,5*]

1 Key Laboratory of Systems Biology, SIBS-Novo Nordisk Translational Research Centre for PreDiabetes, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200233, China
2 Computational Biology Research Center, National Institute of Advanced Industrial Science and Technology, Tokyo 135-0064, Japan
3 INFOCOM Corporation, Tokyo 150-0001, Japan
4 Hefei National Laboratory for Physical Sciences at Microscale and School of Life Sciences, University of Science and Technology of China, Hefei 230027, China
5 Institute of Systems Biology, Shanghai University, Shanghai 200444, China
*corresponding author

Abstract In the aim of identifying significant transcriptional regulatory networks in the liver contributing to diabetes, we have performed comprehensive active regulatory network survey by network screening in 4weeks (w), 8-12w, and 18-20w Goto-Kakizaki (GK) rat liver microarray data. The comprehensive survey of the consistency between the networks and the measured data by the network screening approach in the case of non-insulin dependent diabetes in the GK rat reveals: 1. More pathways are active during inter-middle stage diabetes; 2. Inflammation, hypoxia, increased apoptosis, decreased proliferation, and altered metabolism are characteristics and display as early as 4weeks in GK strain; 3. Diabetes progression accompanies insults and compensations;

Keywords GK rat; Regulatory Network; Network Screening; Active Pathways; Diabetic progression

1 Introduction

Type 2 diabetes mellitus (T2DM) is a complex systemic disease, with significant disorders of metabolism [1]. The liver, a central energy metabolic organ, plays a critical role in the development of diabetes [2]. Although gene expression levels are able to be measured via microarray since 1996, it is difficult to evaluate the contributions of one altered gene expression to a specific disease. One of the reasons is that a whole network picture responsible for a specific phase of diabetes is missing, while a single gene has to be put into a network picture to evaluate its

importance.

In the aim of identifying significant transcriptional regulatory networks in the liver contributing to diabetes, we have performed comprehensive active regulatory network survey by network screening in 4weeks (w), 8-12w, and 18-20w Goto-Kakizaki (GK) rat liver microarray data. We identify active regulatory networks in GK rat by network screening in the following procedure. First, the regulatory networks are compiled by using the known binary relationships between the transcriptional factors and their regulated genes and the biological classification scheme, and second, the consistency of each regulatory network with the microarray data measured in GK rat is estimated to detect the active networks under the corresponding conditions.

2 Materials and Methods

2.1 Network Screening

The candidates of active regulatory networks are detected by network screening in the following manner. First, the regulatory network sets are generated by combining the binary relationships between transcriptional factors (TFs) and their regulating genes, which are compiled in TRANSFAC database [3], and the functional gene sets defined in the Molecular Signatures Database (MSigDB) [4]. Then, we calculate the graph consistency probability (GCP) [5], which expresses the consistency of a given network structure with the monitored expression data of the constituent genes in this study, for each of the network structures obtained at the first step. In addition, in each reference network, the enrichment probability of the genes with the significant differences between GK and WKY rats (expression signature) is further tested. For this purpose, the expression signature is determined using the Student's t-test (for a false discovery rate [FDR] < 5% in expression between GK and WKY rats). The number of genes included in the expression signature is tested for each network, based on the hyper-geometric probability. Thus, we refine the candidates of active regulatory networks, in terms of both the network structure by GCP and the extent of gene expression by enrichment probability. The significance of both thresholds is set to 0.05.

2.2 Microarray data

Microarray dataset is cited from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/projects/geo/) database (GSE 13271). The data are composed of 31,099 probes measured by using Affymetrix Microarray Suite 5.0 (Affymetrix), which are reduced into 14,506 genes, for 5 samples of male Goto-Kakizaki (GK) spontaneously diabetic rats and WKY rats at each of 5 time points (4, 8, 12, 16, and 20 weeks of age). Hyperglycemia begins to show at 4 weeks of age and stabilize after 16 weeks in GK, thus we divided data into three functional groups: 4w, 8-12w, and 16-20w.

3 Results and Discussion

3.1 Activated pathways revealed by network screening and their functions

We identify a total of 20 and 19 differentially activating transcriptional regulatory networks in GK and WKY rats, respectively. There are fewer pathways activating at 4w and 16-20w in GK rats which are at the beginning and the steady state of diabetes. While during 8-12w, more pathways are significantly activated, which indicates a dynamic process involving dysfunctions and compensations in the development of diabetes, as showed outside blood glucose fluctuations. There are more active pathways in the 4w and 8-12w than those in the 16-20w in WKY, which may be due to body growth and development. It is worth pointing out that many activating pathways in WKY are diminished in GK rats at 4w, suggesting that those pathways in the liver important to keep glucose metabolism homeostasis are dysfunction at very early stages of diseases.

Table 1: Active regulatory networks classification according to their functions.

	GK	WKY
Metabolism	HSC_LATEPROGENITORS_ADULT ATRIA_UP GLYCEROPHOSPHOLIPID_METABOLISM GOLUB_ALL_VS_AML_UP HOHENKIRK_MONOCYTE_DEND_UP HSC_LATEPROGENITORS_ADULT LONGEVITYPATHWAY VHL_NORMAL_UP	HASLINGER_B_CLL_MUTATED VEGF_HUVEC_30MIN_UP YAGI_AML_PROG_ASSOC ZHAN_MM_CD138_CD1_VS_REST
Immune	HSC_LATEPROGENITORS_ADULT LINDSTEDT_DEND_8H_VS_48H_DN LEI_HOXC8_DN TESTIS_EXPRESSED_GENES TSADAC_RKOEXP_UP	NGUYEN_KERATO_UP ICF_UP
Transcription	HSC_LATEPROGENITORS_ADULT ATRIA_UP GOLUB_ALL_VS_AML_UP HOHENKIRK_MONOCYTE_DEND_UP HSC_LATEPROGENITORS_ADULT MEF2DPATHWAY P35ALZHEIMERSPATHWAY	VEGFPATHWAY HCC_SURVIVAL_GOOD_VS_POOR_UP HSC_LATEPROGENITORS_SHARED SCHURINGA_STAT5A_UP NUCLEAR_RECEPTORS CELL_DEATH NI2_LUNG_DN PARK_RARALPHA_MOD NUCLEAR_RECEPTORS TGFBPATHWAY
Signaling transduction	INTEGRINPATHWAY INTEGRIN_MEDIATED_CELL_ADHESION_KEGG MEF2DPATHWAY P35ALZHEIMERSPATHWAY RCC_NL_UP VHL_NORMAL_UP ASTON_OLIGODENDROGLIA_MYELINATION_SUBSET BRCA_BRCA1_NEG LEI_HOXC8_DN TESTIS_EXPRESSED_GENES TSADAC_RKOEXP_UP VEGFPATHWAY HSC_LATEPROGENITORS_ADULT	P21_P53_MIDDLE_DN UVB_NHEK1_C2 ALKPATHWAY BRENTANI_PROTEIN_MODIFICATION CELL_DEATH NI2_LUNG_DN PARK_RARALPHA_MOD TGFBPATHWAY NUCLEAR_RECEPTORS

3.2 Classification of activated pathways revealed in terms of their functions

Apart from the view of differentially activated networks along the time points, the networks in the GK and WKY strains can be classified into 4 functional categories in Table 1, which are metabolism, immune, transcription, and signal transduction. Note that some activated pathways share their functions. In that case, they are listed under several functional groups as long as the condition met. Then, we combine the activated networks belonging to the same functional category, if any constituent genes of transcriptional factor (TF) and its regulated gene share each other in the networks. Thus TF-gene expression networks for each functional category are created (Figs 1A-E). Interestingly, significantly activated networks in GK and WKY strains are very different even in the same functional category. We will describe the details of the activated networks in 4 functional categories, below.

3.2.1 Metabolism

Metabolic TF regulatory network in WKY rats reveals increased expression of several genes are important to keep metabolic homeostasis, e.g. bone gamma-carboxyglutamic acid-containing protein (BGLAP), Hepatocyte nuclear factor 4 alpha (HNF4A) and Lipoprotein lipase (LPL) (Fig.1A). In addition to its role in bone-building, BGLAP stimulates pancreatic beta cells releasing more insulin and increases insulin sensitivity via enhancing adipocytes adiponectin secretion [6]. HNF4A plays a key role in liver development. Mutations in this gene have been associated with maturity-onset non-insulin-dependent diabetes of the young (MODY) [7]. Our analysis indicates that reduced HNF4A expression may also favor T2DM development in GK rats. LPL is an enzyme that hydrolyzes triglyceride in lipoproteins such as very low-density lipoproteins (VLDL) and reforms high-density lipoproteins (HDL). Lipoprotein lipase deficiency leads to elevated levels of triglycerides in the bloodstream [8]. Interestingly, like HNF4A, LPL is also suggested to be a diabetes susceptibility gene by human studies [9].

Metabolic networks in GK rats are more complicated than those in WKY rats (Fig.1B). Besides the reduced expression of three genes described in the previous paragraph in diabetic GK rats, some pathways identified by network screening further contribute to metabolism disorders. Cytokines induce activation of the JAK-STAT pathway results in expression of various suppressors of cytokine signaling (SOCS). Expression of SOCS2 and STAT5 but not SOCS3 is decreased in GK rats. Decreased expression of SOCS2 leads to enlarged internal organs, which consists with the description in the original paper that liver weight as a percentage of total body weight is significantly larger in GK [10]. Insulin directly stimulates SOCS2 and STAT5 expression, and the decreased SOCS2 and STAT5 levels are due to insulin deficiency or resistance. IGF-1 (insulin-like growth factor-1) has functions similar to insulin, and it can also improve blood sugar profiles in type 2 diabetics [11]. IGF-1 levels are increased at 4w, but significantly decreased, thereafter may partially explain the insulin resistance after 8 weeks of age in GK rats.

We also observed some compensative pathways activation in GK to fight against

insulin resistance. For instance, insulin receptor substrate 2 (IRS2) is up-regulated and SOCS1 is down-regulated at 8-12w. Cytokine-induced SOCS-1 interacts with the phosphorylated insulin receptor and promotes ubiquitination (Ub) and degradation of IR-IRS complex, thereby preventing insulin signaling pathways [12]. Decreased SOCS-1 is correlated to insulin sensitivity. However, compensations fail to stop development of diabetes.

3.2.2 Immune

Many proinflammatory pathways are activating in GK compared to WKY rats (Fig 1C). From the TF-regulatory gene expression networks in GK rats, two hubs which play important role in immune damages are displayed.

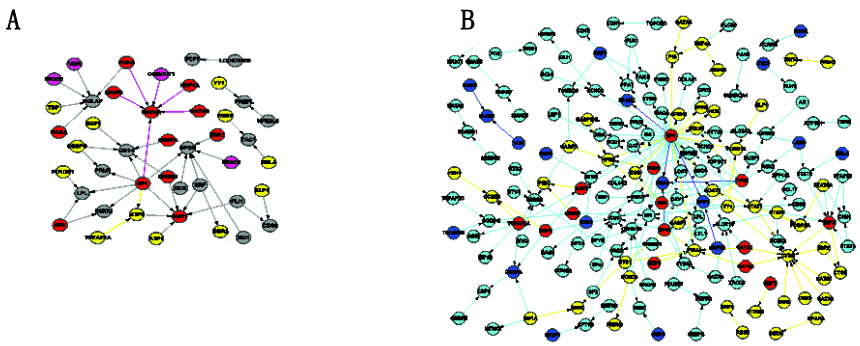
Cytochrome b-245, beta polypeptide (CYBB) is a gene encoding gp91(phox) protein, a phagocyte NADPH oxidase. The protein is also known as P91-PHOX and NOX2. Reactive oxygen species (ROS) produced by NOX2 are able to kill phagocytized bacteria. Because of its highly reactive nature, CYBB has been considered harmful mediators of inflammation [13]. NF-KB and interferon-gamma further increase CYBB expression. Prolonged highly CYBB expression enhanced production of reactive oxygen species, which are critical sources mediating neurovascular damage. Significantly overexpressed CYBB in GK stain is a critical contributor to the microvascular complications associated with diabetes.

Activating transcription factor 3 (ATF3) is a stress-inducible gene and encodes ATF3 transcription factors. ATF3 expression has been reported up-regulated in insulitis and type 1 or type 2 diabetics. Induction of ATF3 is mediated by proinflammatory factors, such as nitric oxide and NF-κB. Importantly, the induction of ATF3 leads cell apoptosis, while signals without ATF3 up-regulation do not cause cell damage [14]. Increased gene expression of ATF3 in GK rats are related to increased immune response and apoptosis.

Besides these two hubs, about 20 immune related genes are changed in GK strain. In sum, inflammation is significantly increased in diabetic Gk rats.

3.2.3 Transcription

Pathways analysis reveals that WKY transcriptional network is a balanced and well-controlled system.



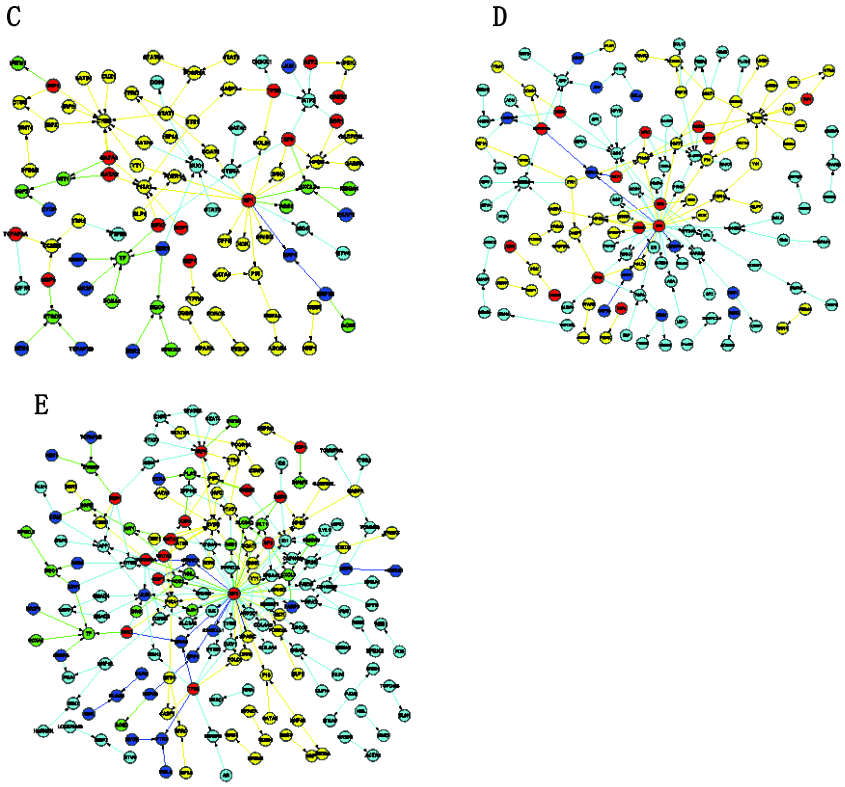


Fig 1. Combined networks in the four functional categories. A: metabolism in WKY, B: metabolism in GK, C: immune in GK, D: transcription in GK, and E: signal transduction in GK (see details in the text).

Some involves in cell replication, good survival and self renewal. Others, including P21-P53_Middle_DN, UBV_NHEK1_C2, and TGFBPATHWAY, emphasize anticancer and cell cycle checkpoints regulation (Table 1).

In GK rats, two out of 7 pathways are related to apoptosis (Table 1 and Fig 1D). Caspase 1 (CASP1), which has been shown to induce cell apoptosis, is overexpressed. Transforming growth factor alpha (TGFA), which stimulates neural cell proliferation, is inhibited. Interestingly, diabetes activates several genes involving in neurodegenerative disorders. Alzheimer's disease shares many commons with T2DM, so that some scientists proposed to call Alzheimer’s disease "type 3 diabetes" or "diabetes of the brain." Calpain small subunit 1 (CAPNS1), a highly-conserved cysteine protease, which have been implicated in neurodegenerative processes after oxidative stress stimulation, is more active in GK. Casein kinase I isoform alpha (CSNK1A1), also called CK1α, is associated with phosphorylate tau and amyloid formation. The expression of CK1α gene is much higher in GK.

3.2.4 Signal Transduction

The key difference in signal transduction category is activation of hypoxia and coagulation related pathways in GK rats (Table 1 and Fig 1E). Coagulation factor XIII A chain (F13A1) is the last zymogen activating in the blood coagulation cascade, which stabilize clots. In GK rats, F13A1 gene expression levels are significantly elevated which enhance thrombosis. Macrophages expressing high affinity immunoglobulin gamma Fc receptor I (FcgRIa) also display coagulation function via binding platelets and initiate thrombosis. Tissue plasminogen activator (PLAT) breakdowns blood clots. GK rats present significantly higher PLAT expression levels, which may explain hemolysis and thrombosis co-existing in diabetics. Dr. Auwerx reported in diabetics, PLAT and plasminogen activator (PA) inhibitor are both activated [16]. The elevated levels of PA-inhibitor activity abolish PLAT activity inducing a reduced fibrinolytic capacity.

3.3 Further remarks

In order to understand the dynamical changes of regulatory networks in the development of diabetes, the active networks can be also interpreted in terms of each time segments. The characteristic features of the active networks, especially relationship between active networks and diabetic progression, will be reported in near future.

This study is the first time to use network screening to explain the role of liver in development of diabetes and the underline mechanism. The results provide many important rational information and insights into guiding experiments design. It is worth pointing out that the molecular relationships change dynamically, depending on the conditions in a living cell, which suggests implicitly that all of the relationships in the knowledge-based network do not always exist.

Acknowledges

We are grateful to Dr. Jiarui Wu, Dr. Jacob Sten Petersen, and Dr. Trine Ryberg Clausen for their comments and support. Work was supported by grants from NN-CAS Research Foundation under NO. NNCAS-2009-1 (H.Z.), and Development of Analysis Technology for Induced Pluripotent Stem (iPS) Cell', from The New Energy and Industrial Technology Development Organization, Japan (S.S. and K.H.).

References

[1] Smyth S, Heron A: Diabetes and obesity: the twin epidemics. *Nat Med* 2006, **12:**75-80.

[2] Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, Boutin P, Vincent D, Belisle A, Hadjadj S, Balkau B, Heude B, Charpentier G, Hudson TJ, Montpetit A, Pshezhetsky AV, Prentki M, Posner BI, Balding DJ, Meyre D, Polychronakos C, Froguel P: A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 2007, **445:**881-885.

[3] Wingender, E: TRANSFAC project as an example of framework technology that supports the analysis of genomic regulation. Brief. Bioinformatics 2008，9:326-332.

[4] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP: Gene set enrichment

analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA 2005, 102:15545-15550.
[5] Saito S, Aburatani S, Horimoto K: Network evaluation from the consistency of the graph structure with the measured data. BMC Sys Biol 2008, 2:84.
[6] Lee NK, Sowa H, Hinoi E, Ferron M, Ahn JD, Confavreux C, Dacquin R, Mee PJ, McKee MD, Jung DY, Zhang Z, Kim JK, Mauvais-Jarvis F, Ducy P, Karsenty G: Endocrine regulation of energy metabolism by the skeleton. Cell 2007, 130:456-469.
[7] Beijers HJ, Losekoot M, Odink RJ, Bravenboer B: Hepatocyte nuclear factor (HNF)1A and HNF4A substitution occurring simultaneously in a family with maturity-onset diabetes of the young. Diabet Med 2009, 26:1172-1174.
[8] Nikkilä EA, Huttunen JK, Ehnholm C: Postheparin plasma lipoprotein lipase and hepatic lipase in diabetes mellitus. Relationship to plasma triglyceride metabolism. Diabetes 1977, 26: 11-21.
[9] Das B, Pawar N, Saini D, Seshadri M: Genetic association study of selected candidate genes (ApoB, LPL, Leptin) and telomere length in obese and hypertensive individuals. BMC Med Genet 2009, 10:99.
[10] Greenhalgh CJ, Rico-Bautista E, Lorentzon M, Thaus AL, Morgan PO, Willson TA, Zervoudakis P, Metcalf D, Street I, Nicola NA, Nash AD, Fabri LJ, Norstedt G, Ohlsson C, Flores-Morales A, Alexander WS, Hilton DJ: SOCS2 negatively regulates growth hormone action in vitro and in vivo. J Clin Invest 2005, 115:397–406.
[11] Haluzik M, Yakar S, Gavrilova O, Setser J, Boisclair Y, LeRoith D: Insulin Resistance in the Liver-Specific IGF-1 Gene-Deleted Mouse Is Abrogated by Deletion of the Acid-Labile Subunit of the IGF-Binding Protein-3 Complex Relative Roles of Growth Hormone and IGF-1 in Insulin Resistance. Diabetes 2003, 52:2483-2489.
[12] Rui L, Yuan M, Frantz D, Shoelson S, White MF: SOCS-1 and SOCS-3 block insulin signaling by ubiquitin-mediated degradation of IRS1 and IRS2. J Biol Chem 2002, 277:42394-42398.
[13] Bolscher BG, de Boer M, de Klein A, Weening RS, Roos D: Point mutations in the beta-subunit of cytochrome b558 leading to X-linked chronic granulomatous disease. Blood 1991, 77: 2482–2487.
[14] Fan F, Jin S, Amundson SA, Tong T, Fan W, Zhao H, Zhu X, Mazzacurati L, Li X, Petrik KL, Fornace AJ Jr, Rajasekaran B, Zhan Q: ATF3 induction following DNA damage is regulated by distinct signaling pathways and over-expression of ATF3 protein suppresses cells growth. Oncogene 2002, 1:7488-7496.
[15] Auwerx J, Bouillon R, Collen D, Geboers J: Tissue-type plasminogen activator antigen and plasminogen activator inhibitor in diabetes mellitus. Arteriosclerosis 1988, 8:68-72.